REMARKS

Status of the claims

With entry of the instant amendment, claims 1-6 have been amended. The amendments add no new matter and are supported throughout the application and claims as filed.

Claims 1-32 are pending in the application. Claims 11-24, 31, and 32 have been withdrawn by the Examiner. Claims 1-10 and 25-30 are therefore under examination.

Objection to the specification

The disclosure was objected to for informalities regarding the sequences shown in Figures 5 and 6. Applicants note that the sequence listing filed July 2, 2004 included SEQ ID NOs for the sequences depicted in Figures 5 and 6. The amendments to the specification presented in this paper provide the SEQ ID NOs. Applicants therefore respectfully request withdrawal of the objection.

Restriction requirement

The Examiner has made the restriction requirement final. Although Applicants disagree with the restriction for reasons of record, claim 1 has been amended to delete the alternative recitation that the polymerase has at least 80% identity to SEQ ID NO:25.

With regard to the Examiner's contention that claim 1 is not a 'genus type' claim (page 3 of the Office Action), Applicants disagree. Claim 1, both before and after amendment, encompasses the species set forth in claim 7 as well as other species that possess the claimed characteristics. Accordingly, it is properly a generic claim. Applicants also note that claim 1 as it relates to the elected reference sequence SEQ ID NO:24 is under examination as a generic claim (*see,* the enablement rejection outlined by the Examiner). Thus, the Examiner's contention that restriction of this claim would still allow the claim to be fully considered on its merits (page 3) appears to relate to the restriction as applied to the reference sequences SEQ ID NO:24 and SEQ ID NO:25. If this interpretation is incorrect, Applicants respectfully request clarification.

With regard to the restriction applied to SEQ ID NO. 2, the polymerase of SEQ ID NO:2 corresponds to the polymerase region of SEQ ID NO:4. Specifically, SEQ ID NO:4 is a fusion polymerase in which the polymerase domain (SEQ ID NO:2) is fused to Sso7d (*see, e.g.,* the last two sentences of paragraph 155, the descriptions of SEQ ID NOs 1-4, and the sequences of SEQ ID NOs 2 and 4). The restriction of SEQ ID NO:2 and SEQ ID NO:4 into separate groups is therefore additionally improper.

The invention

The current claims relates to hybrid polymerases that have residues from more than one parental polymerase protein. Such polymerases are conveniently generated, *e.g.,* by targeting amino acid residues that are not conserved in the parental polymerases. Thus, such a method can involve aligning the parental polymerase amino acid sequences, and at positions that are not identical, introducing either one or the other parental residue that occurs at that position into the polymerase.

Rejection under 35 U.S.C. § 112, first paragraph--enablement

Claims 1-6, 8, 9, 10, 25, and 27-30 were rejected as not properly enabled by the specification. The Examiner alleges that it would require undue experimentation to determine polymerases that fall within the scope of the claims. The Examiner first contends that one of skill in the art would not be able to determine mutations to be introduced into a polymerase that would not adversely affect function. Specifically, the Examiner alleges that the specification provides no guidance as to what residues can be changed. The Examiner also contends that it is not routine in the art to screen for multiple substitutions and that one of skill in the art would expect that any tolerance to the modification of a given protein would diminish with multiple substitutions. Applicants respectfully traverse this rejection.

First, the application in fact provides guidance as to which residues can be substituted in the claimed polymerase sequences. Furthermore, the hybrid polymerases exemplified in the specification demonstrate that more than one substitution can readily occur in

a parental polymerase sequence, *e.g.*, SEQ ID NO:24, without rendering the polymerase non-functional.

In addition, this is a highly advanced art. Extensive structure-functional analyses have been performed on archaeal polymerases. Such information is readily available to one of skill in the art to further determine residues that can reasonably be expected to tolerate substitution.

Thus, as explained below, the examples detailed in the specification, the guidance provided by Applicants for selecting residues for substitution, and the level of knowledge in this advanced art allow one of skill to identify members of the claimed genus of polymerase proteins without undue experimentation.

The application provides guidance

The Examiner alleges that the specification provides no guidance as to sites within the polymerase protein reference sequence(s) that tolerate substitutions. However, this assertion is incorrect. The genus of proteins set forth in claim 1 is derived from two parental polymerases, *Pfu* and Deep Vent®. The reference polymerase sequence SEQ ID NO:24 is the sequence of *Pfu* polymerase. A member of the claimed genus (claim 1) is therefore characterized as follows: (i) it has the specified identity to *Pfu* polymerase; (ii) it comprises the SEQ ID NO:23 signature sequence, which is the nucleotide binding site; and (iii) it comprises at least one substituted residue at a site that is not conserved between *Pfu* and Deep Vent® where the substituted residue is from the Deep Vent® parent polymerase sequence.

In teaching how to make members of this genus, the specification provides an alignment of the *Pfu* and Deep Vent® sequences (Figure 1) where "X" indicates positions at which the sequences vary. The two polymerases differ from one another at 115 locations and have 85% identity over their length (*see, e.g.*, paragraph 144). The specification teaches that hybrid polymerases can be generated based on alternatively incorporating a residue from either one or the other parent polymerase at sites of variation in the sequence(*e.g.*, paragraphs 145-148). The specification also teaches that conservative substitutions can be introduced into amino acid sequences that are reasonably expected to preserve function *(see, e.g.*, paragraphs 59, 60,

and lines 3-13 on page 15). Thus, the specification provides extensive direction to the artisan as to the type and placement of substitutions in a polymerase sequence.

The specification provides examples

In addition to the guidance provided in the specification at the passages noted above, Applicants have taught specific examples of active polymerase proteins that have been obtained by following the teachings in the specification. For example, the Examples section at pages 39 and 40 teaches at least six hybrid polymerase proteins that were isolated from libraries constructed to provide hybrid polymerases having the characteristics set forth in claim 1 (paragraphs 145-150). Four full-length clones were assembled from one set of libraries (designated "Hyb1 to Hyb4"); two full-length clones ("Phy1" and "Phy2") were assembled from a second collection of libraries. These six polymerases were expressed and purified. All of these six hybrid polymerase proteins had DNA polymerase activity (paragraph 154).

Five additional hybrid polymerase clones were isolated from the second library as Sso7d fusions and were designated PhS3 to PhS7. The polymerases were tested for polymerase and exonuclease activity (paragraph 159). Table 1 provides a summary of characteristic of PhS3 to pHS7 and other hybrid polymerases exemplified in the specification. Table 1 shows that the polymerases indicated as PhS1, PhS2, PhS5, Ph7, Hyb1, Hyb2, Hyb3, and HyS4 (SEQ ID NOs. 4, 6, 8, 10, 12, 16, 18, and 20, respectively) all exhibit activity. HyS1 (SEQ ID NO:14), which is also active, is a Sso7d fusion protein comprising the polymerase Hyb1 (SEQ ID NO:12) (*see, e.g.,* Table 1). PhS1 (SEQ ID NO:4) is an Sso7d polymerase fusion comprising a polymerase having the sequence of Phy1 polymerase SEQ ID NO:2.

In summary, the parental polymerase proteins, Deep Vent® and *Pfu* polymerase share at least 85% identity over 700 contiguous amino acids of SEQ ID NO:24. The specification further teaches that residues, *e.g.,* residues that are not conserved in the two parent polymerases, can be altered to incorporate a residue from a different polymerase to derive hybrid polymerases. The claimed polymerase proteins also have the nucleotide binding domain site sequence set forth in SEQ ID NO:23. Applicants have exemplified at least eight active distinct polymerase sequences that were obtained from libraries designed to provide polymerases having

characteristics recited in claim 1. For example, PhS1 has 60 residues that are altered from the *Pfu* parent reference sequence to the Deep Vent® parent residue (see, sequences in Figure 5 for comparison). Other hybrid polymerases, *e.g.*, PhS2, have additional sites that are varied with regard to the parent *Pfu* sequence that are not altered in PhS1. For example, the residues at positions 28 and 30, among others, are unchanged in PhS1 relative to *Pfu*, but are changed in PhS2. The disclosure therefore in fact provides considerable guidance to one of skill as to which residues of the *Pfu* reference sequence can tolerate substitution. In view of the foregoing, one of skill in the art has a reasonable expectation that hybrid polymerase proteins as claimed that retain activity, *e.g.*, either polymerase activity and/or exonuclease activity, can readily be generated in accordance with Applicants' teachings.

The prior art provides detailed structural and functional analyses of polymerase B proteins

In addition to the specific guidance and examples provided in the specification, it is noted that this is an advanced art. The structure of Archaeal family B polymerases is known. For example, Hopfner *et al., Proc. Natl. Acad. Sci USA* 96:2600-2605, 1999 ("Hopfner", attached as Exhibit A) provides a crystal structure of an Archaeal DNA polymerase. Hopfner teaches that the structure of the *Thermococcus gorgonarius (Tgo)* polymerase allows the generation of a structure based sequence alignment of the archaeal subfamily of type B DNA polymerases, including *Pfu* polymerase (p. 3603, column 2, second full paragraph and Figure 3). The authors additionally explain that the location of the polymerase active site and the residues at that site are maintained in the archaeal family (*see, e.g.*, p. 3603, second column, second to fifth full paragraphs). Furthermore, the active site for the 3' to 5' exonuclease was also characterized (p. 3603, second column, sixth full paragraph). Hopfner thus demonstrates that a considerable body of knowledge relating to the structure of archaeal polymerases was available in the art at the time of Applicants' invention. Accordingly, one of skill could also consider this structural information to identify or generate members of the claimed genus that can reasonably be expected to have activity.

The specification teaches one of skill how to practice the claimed invention

As long as the specification discloses at least one method for making and using the claimed invention that bears a *reasonable* correlation to the entire scope of the claim, then the enablement requirement is satisfied (MPEP § 2164.02(b), emphasis added). Applicants' specification meets this requirement.

Applicants have provided a rational scheme for producing the claimed polymerases, as evidenced by the exemplification of multiple functional hybrid polymerases obtained following this scheme. Applicants have provided guidance as to which residues in the parental polymerase proteins can be successfully substituted. Moreover, the advanced state of the art provides the practitioner with additional information regarding polymerase structure and function to further identify regions of the parental polymerase that can be modified with a reasonable expectation of retaining the biological function.

As the Examiner knows, it is not the quantity of experimentation that is required to practice the invention that determines whether the disclosure properly enables the claims, it is whether such experimentation is reasonable. It is well-settled in the biotechnology art that routine screening of even large numbers of samples is not undue experimentation when a probability of success exists. *In re Wands*, 858 F.2d 731, 8 USPQ2d 1400 (Fed. Cir. 1988). As stated in Wands, "enablement is not precluded by the necessity for some experimentation, such as routine screening." *In re Wands*, at 1404. The fact that experimentation may be complex does not render it undue. Applicants have demonstrated that the teachings in the specification, particularly in view of the high level of skill of the ordinary practitioner in this advanced art, provide adequate guidance such that one of skill in the art could practice the invention with a reasonable expectation of success.

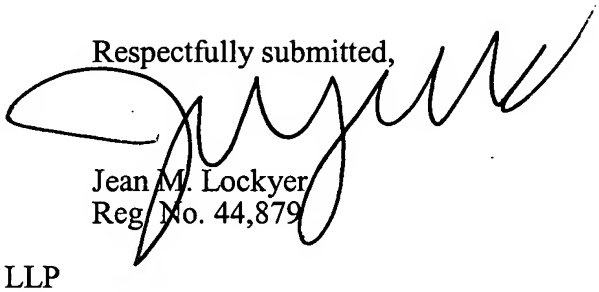
In view of the foregoing, the disclosure properly enables the invention over the scope of the current claims. Applicants therefore respectfully request withdrawal of the rejection.

CONCLUSION

Applicants believe all claims now pending in this Application are in condition for allowance. The issuance of a formal Notice of Allowance at an early date is respectfully requested.

If the Examiner believes a telephone conference would expedite prosecution of this application, please telephone the undersigned at 415-576-0200.

Respectfully submitted,



Jean M. Lockyer
Reg. No. 44,879

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, Eighth Floor
San Francisco, California 94111-3834
Tel: 415-576-0200
Fax: 415-576-0300
JML:jml
60941022 v1

Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 3600–3605, March 1999
Biochemistry

Exhibit A

Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*

(x-ray structure/disulfide bonds/replication/Archaea/exonuclease)

KARL-PETER HOPFNER*†‡, ANDREAS EICHINGER*, RICHARD A. ENGH*§, FRANK LAUE§, WALTRAUD ANKENBAUER§, ROBERT HUBER*, AND BERNHARD ANGERER§

*Abteilung Strukturforschung, Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany; and §Roche Diagnostics, D-82372 Penzberg, Germany

Contributed by Robert Huber, January 22, 1999

ABSTRACT **Most known archaeal DNA polymerases belong to the type B family, which also includes the DNA replication polymerases of eukaryotes, but maintain high fidelity at extreme conditions. We describe here the 2.5 Å resolution crystal structure of a DNA polymerase from the Archaea *Thermococcus gorgonarius* and identify structural features of the fold and the active site that are likely responsible for its thermostable function. Comparison with the mesophilic B type DNA polymerase gp43 of the bacteriophage RB69 highlights thermophilic adaptations, which include the presence of two disulfide bonds and an enhanced electrostatic complementarity at the DNA–protein interface. In contrast to gp43, several loops in the exonuclease and thumb domains are more closely packed; this apparently blocks primer binding to the exonuclease active site. A physiological role of this "closed" conformation is unknown but may represent a polymerase mode, in contrast to an editing mode with an open exonuclease site. This archaeal B DNA polymerase structure provides a starting point for structure-based design of polymerases or ligands with applications in biotechnology and the development of antiviral or anticancer agents.**

Propagation of cells requires faithful DNA replication. This is performed *in vivo* by DNA polymerases (pols), which attach the appropriate dNTP to the nascent DNA primer strand to match its paired template. Different families of pols are involved in different DNA polymerization processes including not only DNA replication (1, 2) but also repair and recombination (3, 4), a heterogeneity also reflected by varying polypeptide structures and/or subunit compositions (3, 5). Some pols complement polymerase activity with 3′ → 5′ exonuclease activity (editing activity) and/or 5′ → 3′ "structure-specific endonuclease" activity, often located in separate structural domains on the same polypeptide chain (4–8).

Crystal structures are available for most known polymerase families, including the A family DNA polymerases (9–14), pol β (15–17), HIV reverse transcriptase (18–20), and recently, the B family pol gp43 from bacteriophage RB69 (21). All share a functional polymerase structure, which resembles a right hand built by the palm, fingers and thumb domains (see ref. 7 for review). Although the fingers and thumb domains are highly diverse among the different families, the palm domains, which contain the conserved catalytic aspartate residues, show a similar topology among all families except pol β. The polymerase nucleotidyl transfer was studied in detail for the A family polymerases, HIV reverse transcriptase, and pol β, and was shown to involve two metal ions (summarized in ref. 7).

Considerably less is known for the family of type B pols, which are replicative enzymes in eukaryotes and most likely also Archaea (22, 23). The structure of gp43 from bacteriophage RB69 (21) provided an excellent first insight into this family. In addition to the three polymerase domains, gp43 contains an 3′ → 5′ exonuclease domain and an N-terminal domain. The exonuclease and palm domains share the topology and active site of A family enzymes, implying similar metal-assisted mechanisms for polymerase and exonuclease activities (21). The thumb and finger domains are apparently unrelated to the other polymerase families. The function of the N-terminal domain remains unknown, but may help assemble the multicomponent replication apparatus (21).

Much is known about the replication of phages (24–26), viruses (1, 27), Prokaryota, (28) and Eukaryota (1, 3, 29, 30), which in general involves pols but also primases, helicases, RNaseH, sliding clamps, and other factors (31). Considerably less is known for archaeal replication, where mostly B type polymerases, similar to eukaryotic replication enzymes pol α and δ, have been identified (6, 22, 23, 32–34). This relative ignorance is surprising, because such crucial biotechnological applications as cloning and PCR require the thermostability and fidelity typical of archaeal polymerases (6). Thus, in addition to satisfying basic research interests, structural information could assist, for example, the engineering of variant enzymes with tailored nucleotide incorporation rates or the design of antiviral and anticancer polymerase inhibitors. For these reasons, we have determined the structure of a DNA polymerase from *Thermococcus gorgonarius* (*Tgo*), an extremely thermophilic sulfur-metabolizing archaeon isolated from a geothermal vent in New Zealand (35). This enzyme possesses pol and a 3′ → 5′ exonuclease activity, which together ensure thermostable replication with high fidelity (error rate: 3.3–2.2 × 10^{-6}; see ref. 36). The 2.5 Å structure shows a topological similarity to gp43 and gives insight in the structural biology of archaeal DNA polymerases, including the identification of several mechanisms for thermophilic adaptation.

MATERIALS AND METHODS

Materials. All materials were of the highest grade commercially available.

Bacterial Strains. *Escherichia coli* LE392 containing pUBS520 was used as described (36). *E. coli* B834 (DE3) (hsd metB) was a generous gift of Nediljko Budisa (Max-Planck-Institut).

Expression Vectors. PBTac2 was obtained from Roche Molecular Biochemicals.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: *Tgo*, *Thermococcus gorgonarius*; pol, DNA polymerase.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, Biology Department, Brookhaven National Laboratory, Upton, NY, 11973 (PDB ID code 1TGO).

†Present address: The Scripps Research Institute, La Jolla, CA 92037.
‡To whom reprint requests should be addressed. email: hopfner@scripps.edu.

Table 1. Data collection and isomorphous replacement

Derivative	Resolution limit, Å	Total observations	Unique observations	Completeness, %	R_{sym}, %	R_{iso}, %	Phasing power
Native	3.0	136,953	21,529	91.9	7.0		
U	3.2	25,981	16,119	89.5	11.9	18.1	0.31
PT1	3.7	65,464	11,692	98.2	12.0	22.6	1.54
PT2	4.0	53,870	6,293	66.9	13.3	13.8	1.77
PT3	3.4	85,226	14,155	92.8	10.5	17.1	1.16
PT4	2.6	387,925	27,487	90.6	5.7	35.5	0.67
PT5	2.7	358,695	24,543	88.7	6.1	36.2	0.83
PTU	3.5	79,629	11,437	82.4	9.4	24.7	1.54
PB	3.5	78,190	11,336	92.9	11.5	12.8	0.41
PBPT	3.4	94,557	14,129	91.6	8.0	19.6	1.04
OS	2.8	600,108	24,026	91.2	5.3	38.4	1.82

Overall figure of merit (15.0–3.0 Å): 0.73

Heavy-atom derivatives were prepared by soaking the crystals in low-salt buffer containing the heavy atom as follows: U, 0.5 mM uranyl acatate for 2 h; PT1, 5 mM K_2PtCl_4 for 1 d, PT2, 5 mM K_2PtCl_6 for 2 d; PT3, saturated *cis*-dichlorodipyridine-Pt(II) for 2 d; PT4, 5 mM K_2PtCl_6 for 7 d; PT5, saturated *cis*-dichlorodipyridine-Pt(II) for 7 d; PTU, PT1 + U; PB, saturated dinitrophenyl-Pb$(NO_3)_2$ for 7 d; PBPT, PB + PT5; OS, saturated K_2OsO_4 for 7 d. NAT1, PT1, PT2, PT3, PTU, PB and PBPT were collected with a Mar imaging plate and U was collected with a Bruker AXS area detector on a Rigaku rotating anode source. All other data sets were collected with a Mar charge-coupled device at beamline BW6 at DESY, Hamburg.

Cloning, Expression, and Protein Purification. The gene for the 89.8-kDa DNA-dependent pol (*Tgo* pol) was cloned from *Tgo* (Deutsche Sammlung von Mikroorganismen no. 8976) and expressed in *E. coli* LE392pUBS520 pBtac2*Tgo* (Deutsche Sammlung von Mikroorganismen no. 11328) as described (36). *Tgo* pol was purified essentially as described (36) with the substitution of the TSK butyl Toyopearl column by Blue-Trisacryl M (Serva) and with an additional concentration step on Poros 50 HQ anion exchange medium (Roche Molecular Biochemicals). Active fractions were combined, concentrated to 30 mg/ml, and transferred to 20 mM sodium phosphate, pH 8.2/10 mM 2-mercaptoethanol/500 mM NaCl.

The gene for a selenomethionine-containing variant of *Tgo* pol (Se-*Tgo* pol) was expressed in *E. coli* B842 (DE3) (hsd metB) using a published protocol (37). Se-*Tgo* pol was purified by using the wild-type protocol.

Crystallization. Crystals of purified *Tgo* pol (or Se-*Tgo* pol) were grown by using sitting-drop vapor-diffusion technique at 4°C with high-salt conditions (2:2 μl protein:reservoir—100 mM Tris, pH 8/2.0M ammonium sulfate) and diffracted to 3.0 Å (in-house) and to 2.5 Å [beamline BW6 at Deutsches Elektronen Synchrotron (DESY), Hamburg]. Low-salt conditions (100 mM Tris, pH 7.0/200 mM ammonium sulfate/30% PEG 400) yielded only poorly diffracting crystals but allowed soaks (including heavy atoms) with some cell constant modulation (a, b, c = 63.6, 105.0, 160.5) but minimal loss of resolution.

Data Collection and Processing. Data were collected with a MAR imaging plate or a Bruker AXS X1000 mounted on a Rigaku rotating anode source, or with a MAR imaging plate or a MAR CCD (charge-coupled device) at beamline BW6 at DESY, Hamburg. The data were processed with SAINT (Bruker AXS), MOSFLM (Mar CCD; ref 38), or DENZO (MAR imaging plate; ref 39), scaled with SCALA (40) or SCALEPACK (39), and reduced with TRUNCATE (40).

Structure Determination. The structure was solved by multiple isomorphous replacement and anomalous scattering (MIRAS) by using data from crystals transferred to low-salt conditions (Table 1). Crystallographic calculations were done with programs from the CCP4 suite (40). Heavy atom positions of major sites were located in difference Patterson maps and were refined with MLPHARE (40) to calculate protein phase angles to 3.5 Å resolution. A partial polyalanine model was built into interpretable portions of secondary structural elements of the MIRAS map by using MAIN (41). The quality of the electron density was improved by phase combination of the partial model with the experimental phases by using SIGMAA (40), and several cycles of solvent flattening to 3.0 Å by using SOLOMON (40). At this stage, no interpretable density was found for a significant portion of the molecule, comprising residues 147–154, 283–306, 653–728, and 752–773.

Model Building and Refinement. The partial model (R factor 35%) was used to phase the 2.5 Å resolution data of the Se-*Tgo* pol (high-salt conditions). The model was oriented with AMORE (40). The correlation coefficient of 22.0% and the R factor of 50.3% showed divergence of the high- and low-salt structures. After bulk solvent correction, anisotropic B factor correction and rigid-body minimization (treating five domains independently), the partial model was iteratively refined and extended with simulated annealing, Powell minimization, restrained individual B factor refinement with CNS (42), and manual model building with MAIN (41) by using data from 25.0–2.5 Å resolution (Table 2, Fig. 1).

RESULTS AND DISCUSSION

Structure of *Tgo* pol. *Tgo* pol is a ring shaped molecule with dimensions 50 Å × 80 Å × 100 Å. The single polypeptide chain of 773 aa is folded into five distinct structural domains (Fig. 2): the N-terminal domain (residues 1–130), the 3′ → 5′ exonuclease domain (131–326), the palm (369–449 and 500–585), fingers (450–499), and thumb (586–773) domains of the polymerase unit, and a helical interdomain insertion (327–368)

Table 2. Crystallographic refinement, high-salt form

Space group	$P2_12_12_1$
Cell dimensions, Å	a = 58.1, b = 105.2, c = 154.2
Observations, 25–2.5 Å	
Total	482,448
Unique	30,451
Completeness, %	
Total	91.1
Last shell	86.0
R_{sym}, %	
Total	7.1
Last shell	26.2
R factor (R_{free}), %	20.9 (27.1)
rms deviation in bond lengths, Å	0.008
rms deviation in bond angles, °	1.5
No. of nonhydrogen atoms	
Protein	6,328
Water	339

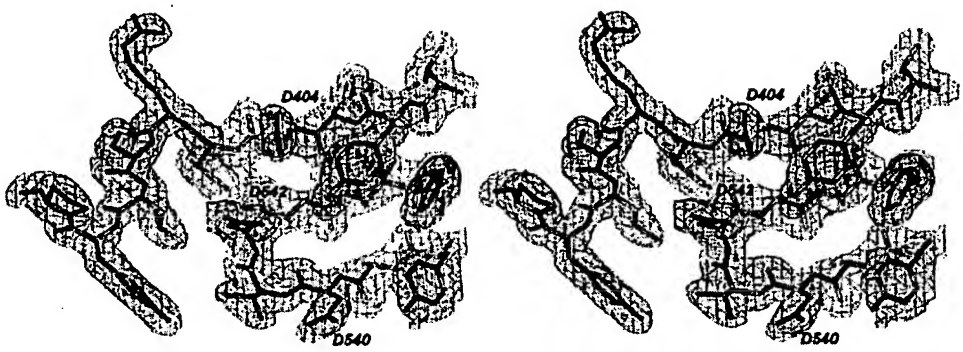


FIG. 1. Stereorepresentation of the electron-density map. The 2 ($F_o - F_c$) electron density contoured at 1.0 σ at the polymerase active site is well defined for the refined model (stick representation).

between the exonuclease and palm domains. The polymerase unit forms the DNA-binding crevice, reminiscent of a right hand, which is the identifying characteristic of pols. Gp43 from bacteriophage RB69 also shows this overall domain topology (21).

Three clefts extend radially from the polymerase active site at the center of the ring: two of them in opposite directions, forming a large cleft across the molecule, and one perpendicular to these. Based on active-site homology to pol A family enzymes, the two opposite clefts probably bind duplex DNA (cleft D, according to ref. 21) and single-strand template DNA (cleft T), respectively. The perpendicular (editing) cleft links the polymerase active site and the exonuclease active site and binds the primer strand in editing mode (21).

The exonuclease domain is structurally equivalent to the 3′ → 5′ exonuclease domain of pol A family (43). Like gp43, however, it is bound at the opposite side of the polymerase unit by noncovalent contacts to the thumb domain at the editing cleft, on one side, and by covalent and noncovalent contacts to the N-terminal and palm domains and the 42 residue interdomain helix, on the other side. This latter segment is located at the base of cleft T, which is additionally bounded by the exonuclease, N-terminal, and palm domains.

The topology of the palm domain is conserved among polymerase families (5), with two long helices (Q and R)

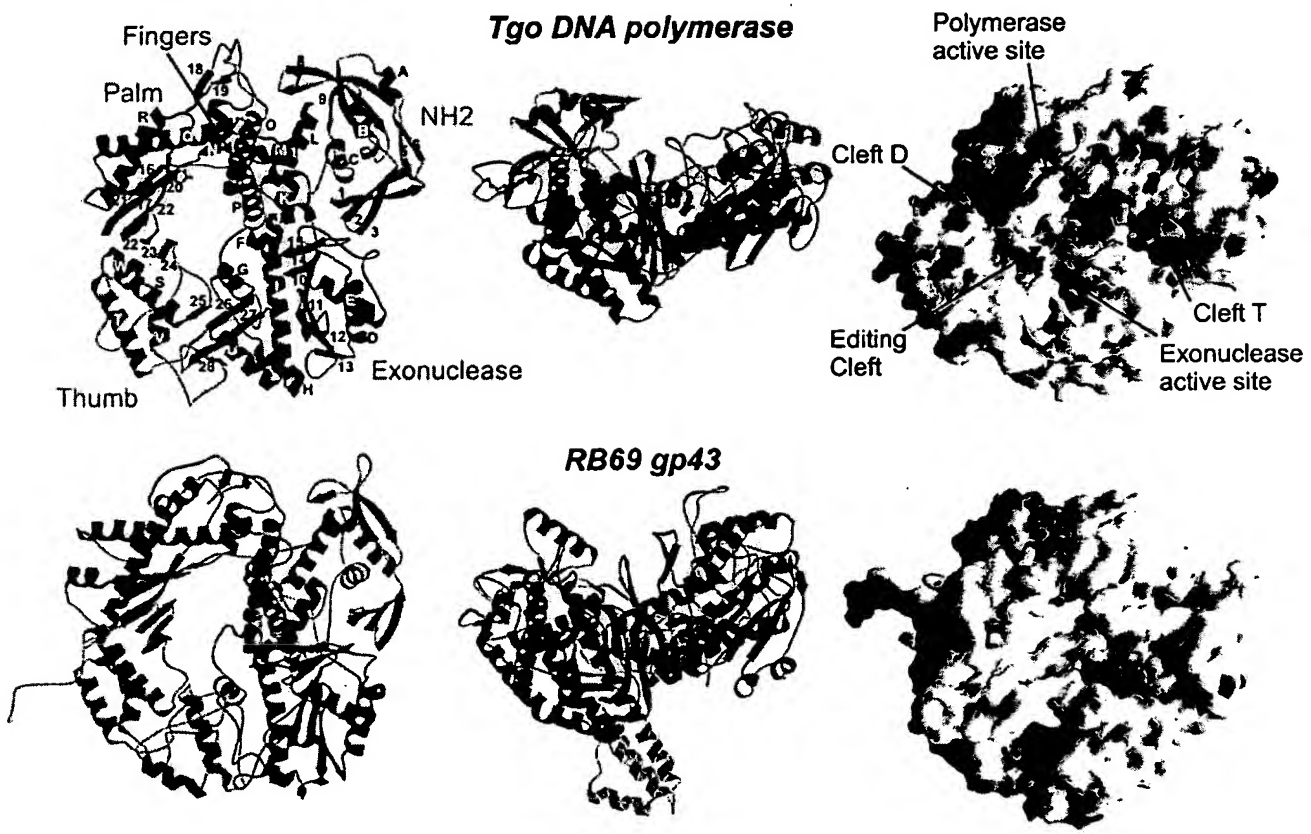


FIG. 2. Structure of *Tgo* pol and comparison with gp43 form bacteriophage RB69. (*Left* and *middle*) Ribbon representation of *Tgo* DNA polymerase (*Upper*) in two orthogonal views with labeled secondary structure elements. The molecule is composed of five domains: N-terminal domain (yellow), 3′ → 5′ exonuclease (red), palm [light and dark magenta represent the N-terminal and C-terminal segment, respectively (see text)], fingers (blue), and thumb (green), which are arranged to form a ring. An interdomain helical segment between the exonuclease domain and the palm is orange. The conserved carboxylates in the active site and the two disulfide bridges are shown as magenta and yellow ball-and-sticks, respectively. *Tgo* pol has the same overall architecture and domain topology than gp43 of RB69 (*Lower*). The 50-residue insertion in the fingers of gp43 is gray. (*Right*) Comparison of molecular surfaces of *Tgo* pol (*Upper*) and RB69 gp43 (*Lower*). Red and blue denote negative and positive electrostatic surface potentials, respectively. In contrast to gp43, *Tgo* pol has a strongly enhanced positive potential at the putative DNA-binding clefts.

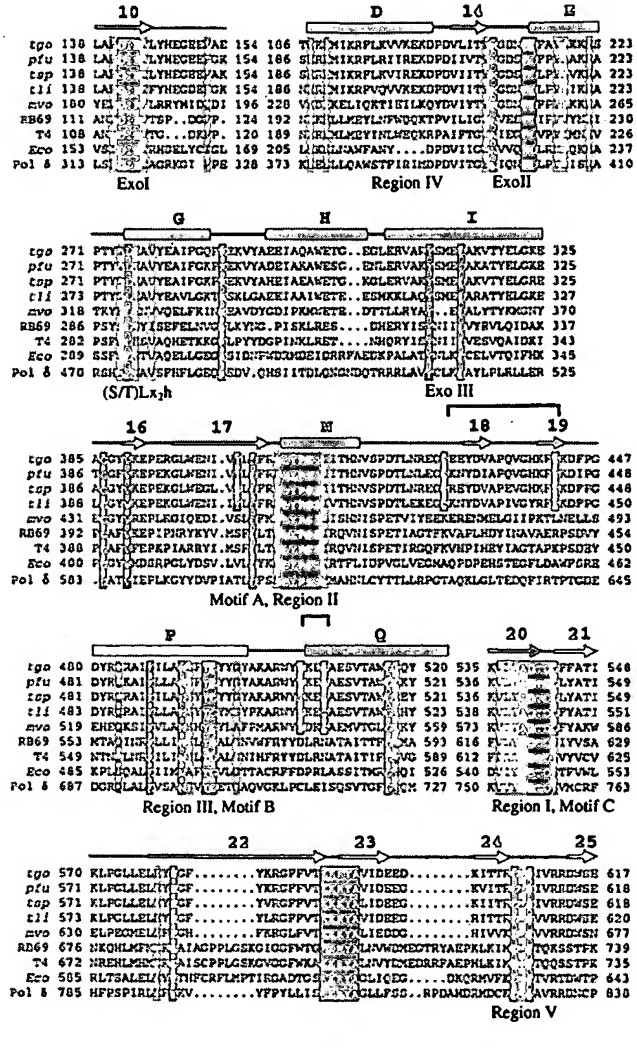


FIG. 3. Sequence alignment of B family DNA polymerases. The alignment has been adapted from ref. 21 to highlight specific residues from the class of archaeal pols. The secondary structure of *Tgo* pol is indicated on top of the alignment with helices (bars), strands (arrows) and loops (lines) colored according to domains with the same color code as Fig. 2*B*. Strictly conserved residues of type B polymerases are red, and additional conserved residues are yellow. Uniquely conserved residues of archaeal type B enzymes—as discussed in the text—are green. Disulfide bonds are shown by a bar on top of the alignment. Abbreviations: *tgo*, *Thermococcus gorgonarius* pol; *pfu*, *Pyrococcus furiosus* pol; *tsp*, *Thermococcus* sp. pol; *tli*, *Thermococcus litoralis* pol; *mvo*, *Methanococcus voltae* pol; RB69, bacteriophage RB69 pol; T4, bacteriophage T4 pol; *Eco*; *E. coli* pol II; pol δ; human pol δ.

packed against the five-stranded antiparallel β-sheet that contains the three conserved aspartate residues involved in nucleotidyl transfer. The fingers emerge from the palm domain as an α-helix-rich insertion. Its 50 residues are folded into two antiparallel coiled α-helices of approximately equal size: helix P contains the conserved $Kx_3NsxYGx_2G$ motif of B type polymerases and is related to the O helix of A type enzymes (see below). The ≈50-residue insertion between helix O and P in RB69 and T4 gp43 is missing in *Tgo* pol, where both helices and 4 residue linker are much shorter than their equivalents in gp43. The shorter fingers of *Tgo* pol presumably reflect the typical structure of the nonbacteriophage B type fingers (pol α, pol δ, and *E. coli* pol II). The thumb domain topology, similar to that of gp69, is unrelated to other polymerase types. However, like the thumb of A type enzymes, a bundle of α-helices at its base protrude from the active site β-sheet. Distal to the active site, the thumb contains a 75-residue subdomain (665–729), which fixes the exonuclease domain and contributes to the editing channel, explaining why mutations in the exonuclease domain of B-type polymerases affect the polymerase activity and vice versa (44; 45).

Weakly defined density across the base of the thumb domain was modeled as the C-terminal 6 residues with a polyalanine chain. The C terminus thus does not protrude from the core molecule as in the RB69 polymerase (21). Because the C terminus of the T4 pol are involved in sliding-clamp binding (46), it is likely, however, that these residues become ordered on any similar holoenzyme formation.

Sequence Alignment of Archaeal DNA Polymerases. The structure of *Tgo* pol allows the generation of a structure based sequence alignment of the archaeal subfamily of type B DNA polymerases, the location of conserved and unique residues, and the comparison with other type B DNA polymerases (Fig. 3).

Polymerase Active Site. The polymerase active site is formed by the central β-sheet (strands 16, 17, 20, 21, and 22) and helix N of the palm domain and helix P located in the fingers and is highly conserved among B family polymerases (Fig. 4). Three carboxylates required for nucleotidyl transfer in B family polymerases, two of which coordinate two metal ions (14) are superimposably conserved among A family enzymes, B family enzymes, and reverse transcriptase (21). Superposition of *Tgo* pol and T7 replication complex (14) places the dNTP near the proposed nucleotide-binding site in helix P, the $K487x_3NSxYGx_2G$ motif (Fig. 3) and suggests interactions of the carboxylates with metals and the phosphate tail of the bound dNTP (Fig. 4). Reorientation of the strictly conserved Lys-487 allows it to mimic the Lys-522–phosphate tail interaction in T7. Tyr-494 ($Kx_3NSxY494Gx_2G$) and Tyr-409 (SLY409PSII) form the bottom of the nucleotide-binding site.

The active site of B family polymerases contains a DTDS motif, which, however is DTDG in the archaeal subfamily. In *Tgo* pol, the relatively conserved Tyr-402 from the adjacent strand provides an alternate alcohol group, at a position appropriate for metal coordination or binding of the 3′ end of the primer. The orientation of Tyr-402 is stabilized by an aromatic cluster that also includes Phe-545 and Tyr-538. Archaeal *Methanococcus voltae* and *Thermococcus* sp. pol's (see Fig. 3) have Tyr at position 545—Phe in *Tgo*—rather than at 402, but might also supply an alcohol group. The displacement of a functional alcohol from serine in DTDS to Tyr-402 or Tyr-545 might stabilize its orientation as an adaptation for thermostability.

The conserved cluster of acidic amino acids (E578, E580) form an unexpected metal-binding site for Mn^{2+} and Zn^{2+} (Fig. 4). Its proximity to Asp-404 and to the expected location of the dNTP γ-phosphate suggests a supporting role in nucleotide binding and/or catalysis.

3′ → 5′ Exonuclease Active Site. *Tgo* pol is characterized by a strong 3′ → 5′ exonuclease activity, unlike eukaryotic B type polymerases (unpublished results). The exonuclease active site is formed at the interface between the exonuclease domain and the tip of the thumb (Fig. 5). All residues required for catalysis are located in the exonuclease domain, which, at least for T4 gp43, retains activity when dissociated from the polymerase (47). However, the thumb domain, with, for example, RB69 gp43's Phe-123–base intercalation, partially controls the binding geometry of single strand DNA (21, 43).

The exonuclease structures of *Tgo* and gp43 DNA polymerases are similar at the editing site but differ considerably at the exonuclease–thumb interface. Strand 10 contains the metal-binding D141IE motif and readily superimposes with the equivalent strand from gp43, allowing modeling of a single-strand DNA segment into the exonuclease site based on theRB69 gp43-$p(dT)_4$ complexes (21). The conserved residues Asp-141 and Glu-143 in the Exo I motif, Tyr-209, Asn-210, Phe-214, and Asp-215 in Exo II, and Tyr-311 and Asp-315 in Exo III are in approximate DNA-binding conformations (Fig.

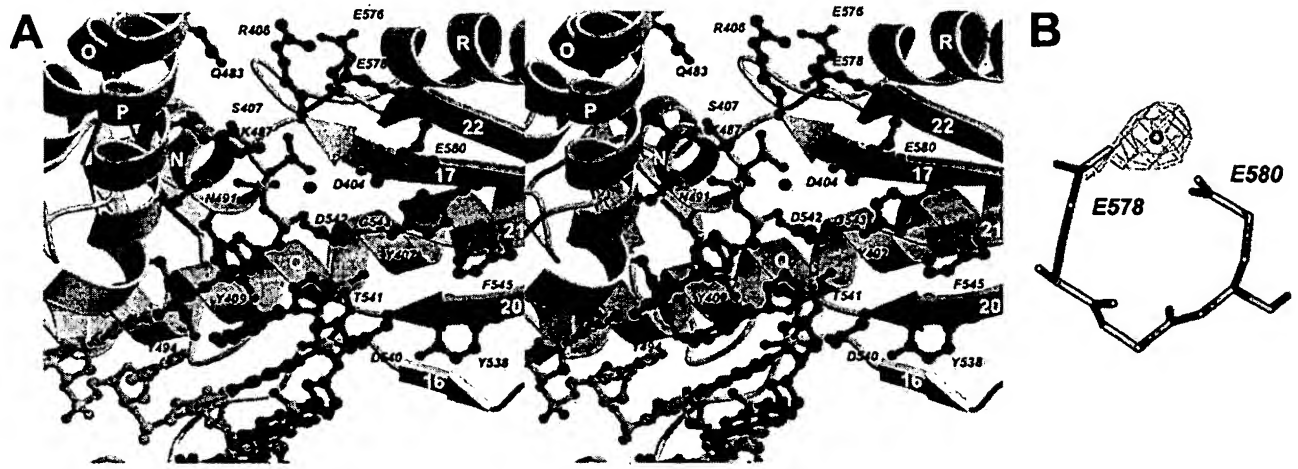


FIG. 4. Polymerase active site. (*A*) Stereoribbon representation (using color code as in Fig. 2) with modeled DNA. Active-site residues are shown as ball-and-stick representations with carbon (green), nitrogen (blue), and oxygen (red) atoms. The DNA template (light brown), primer (light brown), and dNTP (orange) complex has been taken from the coordinates of T7 replication complex (15) by superimposing D404, D542, and adjacent residues with corresponding residues in T7 pol (D475 and D654). Phosphorus atoms are yellow. The two metals of the T7 replication complex are shown as magenta spheres. (*B*) Experimentally observed metal-binding site for Mn^{2+} ($F_o - F_c$ density contoured at 5σ) and Zn^{2+} in the "low salt" crystal form. The carboxylates E578 and E580 are conserved in type B polymerases (Fig. 3).

5*A*). However, the editing cleft is constricted by a displacement of the tip of the thumb toward the exonuclease domain to prohibit single-strand binding (Fig. 5*B*). This shift is correlated with a large change in the loop between strands 10 and 11. In RB69 gp43 (and likewise in T4 gp43), this loop forms a lid over the 3′ base and contains Phe-123, which intercalates between the first two bases. In *Tgo* pol, the loop is curved outward, away from the thumb, and Phe-152 (the equivalent of gp43's Phe-123) attaches to Phe-214 10 Å away from intercalation site. This shift allows the tip of the thumb to move into the editing channel and to block the exonuclease site.

Are There Different Conformations in Polymerase and Editing Mode? If a closed conformation of the exonuclease domain prohibits single strand binding, an open conformation is required for editing. The observed closed conformation may represents the enzyme in "polymerase" mode. Preliminary analysis of the crystal structure of *Tgo* pol in the low-salt conditions indicates a structural change at the interface of exonuclease and thumb, possibly reflecting a transition between open and closed forms. The closed conformation observed here may, however, be a nonphysiological artifact of the high ionic strength used for crystallization. Crystal structures of the enzyme in both polymerase and editing modes are required.

Adaptation to High Temperatures. *Tgo* is a sulfur-metabolizing, extremely thermophilic archaeon, with a growth range between 55°C and 98°C. For accurate replication at this temperature range, the polymerase must not only be stable, but must also adequately bind substrate DNA. A comparison with gp43 from the mesophilic bacteriophage RB69 indicates several such adaptations to high temperatures. Several loops are shorter in *Tgo* pol than in gp43 (Fig. 2), and there is an increase in hydrogen bonded β-strand content: *Tgo* pol secondary structure includes 41% helix, 22% β-strands, and 19% turns (calculated according to ref. 48), whereas gp43 has 42% helix, 17% β-strands, and 19% turns.

Although rare among cytoplasmic or nuclear proteins, two disulfide bridges might be formed: cysteine pairs 428–442 and

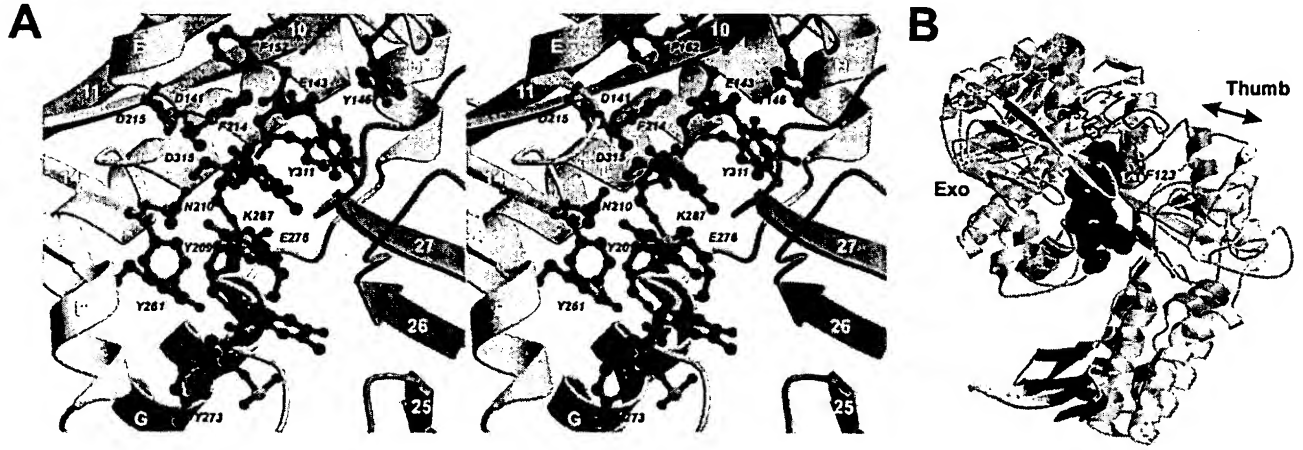


FIG. 5. 3′ → 5′ exonuclease active site. (*A*) Stereoribbon representation with modeled DNA using the color code of Fig. 4*A* (ball-and-stick) and Fig. 2 (ribbons). Active-site residues are shown as ball-and-stick representation. The single-stranded DNA has been taken from the coordinates of RB69–single-stranded DNA complex (21). The orientation of the DNA has been obtained by superimposing the DIE143 motif in *Tgo* pol with corresponding motif of RB69 gp43 (DIE116). Strand 27 and its preceding loop from the thumb (green) is apparently in collision with the modeled DNA. (*B*) Comparison of the exonuclease–thumb interface between *Tgo* pol (color code of Fig. 2*B*) and RB69 gp43 (gray). In *Tgo* pol, the lid of the editing site (red) is bent outward compared with the equivalent loop of gp43 (yellow), allowing the tip of the thumb to move several Å (↘) closer to the exonuclease. This conformation is incompatible with formation of an editing complex [the $p(dT)_4$ of gp43 is shown as brown space-filling model].

506–509, although reduced, are poised for attachment (Fig. 2). Model refinement and electron density inspection with and without constraints for the disulfide bonds verified the reduced state (observed unrestrained SG–SG distance: 2.8 Å and 3.0 Å). This is consistent with our *E. coli* expression and further rules out structural perturbation by nonnative oxidation. These cysteines are located in the palm domain and are conserved among B type enzymes from hyperthermophilic sulfur-metabolizing archaeons, but not among mesophile homologs (Fig. 3). The Cys-428–Cys-442 bridge stabilizes the compact fold of the loop segment between helix N in the palm domain and helix O in the fingers and presumably also the relative orientation of these helices. In addition, the loop segment packs against helix Q in the palm domain. Helix Q, the spine of the palm domain, is further stabilized at the first helical turn by the second disulfide bridge between Cys-506 and Cys-509.

A much enhanced complementary positive potential for all three DNA-binding clefts of *Tgo* pol is observed relative to gp43 (Fig. 2). Thus, in addition to hydrogen bonding and specific DNA–protein interactions, binding to *Tgo* pol has an additional strong stabilizing electrostatic component.

An increase in the number of salt bridges is often associated with thermostability. Although *Tgo* pol has a greater total number of charged residues (262) than gp43 (245), both molecules have 54 salt bridges within a 3–5 Å bound. However, in the 5–7 Å range of charge distance, *Tgo* pol has 77 ion pairs compared with 43 for gp43. This large increase results in a more highly charged surface of *Tgo* pol, accompanied by a more balanced charge distribution, compared with gp43 where charges are often located in patches (Fig. 2).

We thank Dr. Uwe Jacob and Martin Augustin for stimulating discussions, Dr. Nediljko Budisa for help with expression of selenomethionine variants, and Dr. Hans Bartunik and Dr. Gleb Bourenkov at DESY, Hamburg for help with data collection.

1. Waga, S. & Stillman, B. (1998) *Annu. Rev. Biochem.* **67,** 721–751.
2. Maga, G. & Hübscher, U. (1996) *Biochemistry* **35,** 5764–5777.
3. Wang, T. S. F. (1991) *Annu. Rev. Biochem.* **60,** 513–552.
4. Kornberg, A. & Baker, T. A. (1992) *DNA Replication* (Freeman, New York).
5. Joyce, C. M. & Steitz, T. A. (1994) *Annu. Rev. Biochem.* **63,** 777–822.
6. Perler, F. B., Kumar, S. & Kong, H. (1996) *Adv. Protein Chem.* **48,** 377–435.
7. Brautigam, C. A. & Steitz, T. A. (1998) *Curr. Opin. Struct. Biol.* **8,** 54–63.
8. Arnold, E., Ding, J., Hughes, S. H. & Hostomsky, Z. (1995) *Curr. Opin. Struct. Biol.* **5,** 27–38.
9. Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz, T. A. (1985) *Nature (London)* **313,** 762–766.
10. Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R. & Steitz, T. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8924–8928.
11. Kiefer, J. R., Mao, C., Hansen, C. J., Basehore, S. L., Hogrefe, H. H., Braman, J. C. & Beese, L. S. (1997) *Structure (London)* **5,** 95–108.
12. Kiefer, J. R., Mao, C., Braman, J. C. & Beese, L. S. (1998) *Nature (London)* **391,** 304–307.
13. Kim, Y., Eom, S. H., Wang, J., Lee, D. S., Suh, S. W. & Steitz, T. A. (1995) *Nature (London)* **376,** 612–616.
14. Doublie, S., Tabor, S., Long, A. M., Richardson, C. C. & Ellenberger, T. (1998) *Nature (London)* **391,** 251–258.
15. Davies, J. F., Almassy, R. J., Hostomska, Z., Ferre, R. A. & Hostomsky, Z. (1994) *Cell* **76,** 1123–1133.
16. Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H. & Kraut, J. (1994) *Science* **264,** 1930–1935.
17. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H. & Kraut, J. (1994) *Science* **264,** 1891–1903.
18. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. & Steitz, T. A. (1992) *Science* **256,** 1783–1790.
19. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark P., *et al.* (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6320–6324.
20. Jager, J., Smerdon, S. J., Wang, J., Boisvert, D. C. & Steitz, T. A. (1994) *Structure (London)* **2,** 869–876.
21. Wang, J., Sattar, A. K., Wang, C. C., Karam, J. D., Konigsberg, W. H. & Steitz, T. A. (1997) *Cell* **89,** 1087–1099.
22. Edgell, D. R. & Doolittle, W. F. (1997) *Cell* **89,** 995–998.
23. Cann, I. K. O., Komori, K., Toh, H., Kanai, S. & Ishino, Y. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14250–14255.
24. Debyser, Z., Tabor, S. & Richardson, C. C. (1994) *Cell* **77,** 157–166.
25. Nossal, N. G. (1994) *Molecular Biology of Bacteriophage T4,* ed. Karam, J. D. (Am. Soc. Microbiol., Washington, DC), pp. 43–53.
26. Nossal, N. G. (1992) *FASEB J.* **6**(3), 871–878.
27. Fanning, E. & Knippers, R. (1992) *Annu. Rev. Biochem.* **61,** 55–85.
28. Kelman, Z. & O'Donnell, M. (1996) *Annu. Rev. Biochem.* **64,** 171–200.
29. De Pamphilis, M. L., ed. (1996) *DNA Replication in Eukaryotic Cells* (Cold Spring Harbor Lab. Press, Plainview, NY).
30. Waga, S. & Stillman, B. (1994) *Nature (London)* **369,** 207–212.
31. Baker, T. A. & Bell, S. P. (1998) *Cell* **92,** 295–305.
32. Uemori, T., Sato, Y., Kato, I., Doi, H. & Ishino, Y. (1997) *Genes Cells* **2**(8), 499–512.
33. Braithwaite, D. K. & Ito, J. (1993) *Nucleic Acids Res.* **21**(4), 787–802.
34. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., Fitzgerald, L. M., Clayton, R. A., Gocayre, J. D. & Venter, J. C. (1996) *Science* **273,** 1058–1073.
35. Miroshnichenko, M. L., Gongadze, G. M., Rainey, F. A., Kostyukova A. S., Lysenko, A. M., Chernyh, N. A. & Bonch-Osmolovskaya, E. A. (1998) *Int. J. Syst. Bacteriol.* **48,** 23–29.
36. Bonch-Osmolovskaya, E., Svetlichny, V., Ankenbauer, W., Schmitz-Agheguian, G., Angerer, B., Ebenbichler, C. & Laue, F. (1996) Eur. Patent Appl. EP 0 834 570 A1.
37. Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. & Huber, R. (1995) *Eur. J. Biochem.* **320,** 788–796.
38. Leslie, A. G. W. (1994) MOSFLM User Guide (MRC Laboratory of Molecular Biology, Cambridge, UK), Version 5.1.
39. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
40. Collaborative Computational Project, No. 4 (1994) *Acta Crystallogr. D* **50,** 760–763.
41. Turk, D. (1992) Ph.D. Thesis (Technische Universität, München).
42. Adams, P. D., Pannu, N. S., Read, R. J. & Brunger, A. T. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 5018–5023.
43. Wang, J., Yu, P., Lin, T. C. Konigsberg, W. H. & Steitz, T. A. (1996) *Biochemistry* **35,** 8110–8119.
44. Sattar A. K., Lin, T. C., Jones, C. & Konigsberg, W. H. (1996) *Biochemistry* **35,** 16621–16629.
45. Reha-Krantz, L. J. & Nonay, R. L. (1994) *J. Biol. Chem.* **269,** 5635–5643.
46. Berdis, A. J., Soumillion, P. & Benkovic, S. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 12822–12827.
47. Lin, T. C., Karam, G. & Konigsberg, W. H. (1994) *J. Biol. Chem.* **269,** 19286–19394.
48. Kabsch, W. & Sander C. (1983) *Biopolymers* **22,** 2577–2637.